Ph.D. Title Proposal

On

"Enhancing Data Efficiency in Visual Classification through a Generative and Few-Shot Learning Framework"

Submitted in the partial fulfilment of the requirement

for the award of the Degree of

Doctor of Philosophy in Computer Applications

By

Prashant Ashok Chaubey

Under the Guidance of

Dr. Priya Chandran

Computer Applications

Bharati Vidyapeeth's

Institute of Management and Information Technology

CBD Belapur, Navi Mumbai -400614

University of Mumbai

PROPOSAL FOR TOPIC APPROVAL TO BE SUBMITTED TO THE UNIVERSITY OF MUMBAI FOR THE PH.D. DEGREE IN THE SUBJECT OF COMPUTER APPLICATIONS

Title of the topic	:	Enhancing Data Efficiency in Visual Classification through a Generative and Few-Shot Learning Framework
Name of the student	:	PRASHANT ASHOK CHAUBEY
Basic Qualification	:	MCA
In the case of Ph.D. student Year of Passing of PET/NET/SET/M.Phil /GPAT/GATE (Teaching Experience)	:	June -2019
Name of the Research Supervisor/Guide	:	Dr. Priya Chandran
Date of Superannuation/Retirement of the Research Supervisor/Guide	:	31/5/2038
Name of the Research Centre	:	Bharati Vidyapeeth's Institute of Management and Information Technology, CBD Belapur, Navi Mumbai
Signature of the Student	:	
Signature of the Research Supervisor/Guide	:	

Date: 14/8/25

B. V.'s Institute of Management & Information Technology NAVI MUMBAI

Signature of Principal

PRINCIPAL
Bharati Vidyapeeth's
Institute of Management & Information Technology
Sector-8, C.B.D. Belapur,
Navi Mumbai-400 614

Declaration

I wish to state that the work embodied in this synopsis titled “Enhancing Data Efficiency in Visual Classification through a Generative and Few-Shot Learning Framework” forms my own contribution to the work carried out under the guidance of Dr. Priya Chandran at the Bharti Vidyapeeth’s Institute of Management and Information Technology. I declare that this written submission represents my ideas in my own words and where others’ ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission.

Prashant Ashok Chaubey

Table of Contents

Sr. No	Topic Name	Page No.
	Abstract	1
1	Introduction	2
2	A Survey of Work Done in Research Area	3
3	Gap in the Literature Survey	5
4	Rationale and Significance of Study	7
5	Aim and Objectives of Study	8
6	Methodologies and Techniques to be Used	9
7	Plan of Research	14
8	Expected Outcome of Study	16
9	References	17

Enhancing Data Efficiency in Visual Classification through Generative and Few-Shot Learning Framework

Abstract

Few-shot learning (FSL) aims to enable models to learn new tasks from a minimal number of labelled examples. This is a critical capability in domains where large-scale data collection is expensive, time-consuming, or infeasible. Generative models, particularly diffusion models, offer a promising solution by creating synthetic data to supplement small training sets. However, most existing approaches use pre-trained generators in an offline mode, which prevents adaptation to task-specific requirements.

Recent studies exploring joint training of generators and classifiers show potential for adaptive augmentation, but they face key limitations: instability in training, high computational overhead, misalignment between generative quality metrics and classification utility, and reduced robustness under domain shift or rare-class scenarios.

This research proposes a stable and efficient framework for joint generator–classifier training in few-shot learning. The approach aims to produce synthetic data that directly improves classifier performance, reduce training instability, and introduce evaluation strategies that focus on real task utility. The expected outcome is a practical and adaptable solution that bridges the gap between generative model quality and few-shot classification performance.

1 Introduction

Few-shot learning has emerged as an important research area in machine learning, driven by the need to train models effectively with very limited labelled data. Many practical applications — such as disease detection in medical imaging, fault diagnosis in manufacturing, and species identification in ecology — cannot rely on large, well-annotated datasets due to the cost, time, or expertise required to collect them.

Early approaches to few-shot learning focused on metric-based methods, which learn feature spaces that allow effective classification using similarity measures. Optimization-based strategies followed, enabling models to adapt quickly to new tasks through meta-learned initialization. While these methods achieved strong results on benchmark datasets, their reliance on the limited support set constrained their ability to generalize in complex, real-world scenarios.

The integration of generative models brought new possibilities, allowing artificial expansion of the support set to increase diversity and improve generalization. GANs and VAEs were the first widely used generative techniques in this context, but diffusion models have since surpassed them in producing high-quality, diverse synthetic samples. Despite this progress, most diffusion-based augmentation strategies remain offline, limiting their adaptability to specific tasks.

Recent research has explored the joint training of a generator and classifier in an online, task-adaptive setting, with the generator producing data tailored to the classifier's needs in real time. While promising, such approaches face challenges including unstable training dynamics, high computational costs, and evaluation methods that fail to directly capture their effect on classification accuracy.

This proposal addresses these shortcomings by developing a stable, efficient, and utility-focused joint training framework for few-shot learning. The following sections provide a detailed review of related work, identify the key gaps in existing research, and present the rationale for pursuing this approach.

2 Survey of Work Done in the Research Area

Few-shot learning (FSL) focuses on enabling models to generalize to new tasks with very limited libelled data. This problem is common in domains such as medical diagnosis, environmental monitoring, and remote sensing, where collecting large annotated datasets is costly, time-consuming, or impractical. Over the years, research in FSL has progressed from embedding-based classification methods to optimization strategies and, more recently, to augmentation approaches using generative models, including diffusion-based methods.

2.1 Metric-Based Few-Shot Learning

Metric-based methods aim to learn a feature embedding space in which intra-class distances are small and inter-class distances are large. Early examples follow this paradigm by computing similarity between query and support samples in a learned metric space. Other approaches introduce a learnable comparator to improve similarity computation. While effective in benchmark evaluations, these methods rely entirely on the limited support set and do not introduce new data variations to enhance generalization.

2.2 Optimization-Based Few-Shot Learning

Optimization-based approaches focus on learning model parameters that can be rapidly adapted to new tasks with a few gradient updates. This enables fast adaptation through a meta-learned initialization and has inspired several variations to improve adaptation efficiency. However, these approaches are sensitive to task distribution shifts and do not inherently solve the problem of limited diversity in the support set.

2.3 Data Augmentation in FSL

To improve generalization, researchers have explored expanding the support set through data augmentation. Early work used geometric and photometric transformations, while later studies employed generative models such as GANs and VAEs to create synthetic samples. Although these generative approaches can increase data diversity, GANs and VAEs often produce images lacking in fine detail, which can limit their benefit in high-variance classes.

2.4 Diffusion Models for Image Generation and FSL

Diffusion models have recently emerged as a powerful alternative to GANs and VAEs, capable of generating diverse, high-fidelity samples. In FSL, they have mainly been applied in **offline augmentation** settings, where the generator is pre-trained and then used to create synthetic data prior to meta-learning. This reduces computational demands during training but limits the generator's ability to adapt to the specific task at hand, potentially reducing its impact on classification performance.

2.5 Joint Generator–Classifier Training

Some recent studies have explored training the generator and classifier together in an episodic meta-learning loop, enabling the generator to adapt its outputs for the classifier in real time. While this approach can improve task-specific performance, it introduces significant computational overhead, instability in joint optimization, and limited evaluation beyond perceptual quality metrics. Moreover, many methods still show reduced performance when applied to domain-shifted or rare-class scenarios.

2.6 Summary

Research in few-shot learning has progressed from metric-based and optimization-based methods toward augmentation-driven strategies using advanced generative models. Diffusion models currently set the standard for sample quality but are mostly used in offline settings. The relatively new concept of online, task-adaptive joint training offers adaptability but faces computational, stability, and evaluation challenges. These trends set the stage for the next section, which examines the unresolved challenges in depth.

3 Gap in The Literature Survey

The review of existing literature highlights the significant progress made in few-shot learning (FSL), particularly with the integration of generative models to enhance limited datasets. Despite these advancements, several gaps remain unaddressed, limiting the practical adoption and effectiveness of current approaches in real-world scenarios.

3.1 Lack of Task-Adaptive Generation

Most diffusion-based augmentation methods operate in an offline setting, where the generator is pre-trained on a large dataset and then fixed during few-shot learning. While this improves training efficiency, it prevents the generator from adapting to the specific characteristics of each target task. As a result, the generated data may not optimally address the classifier's needs in different contexts, reducing the potential benefit of augmentation.

3.2 Instability in Joint Training

A smaller body of research has investigated online, task-adaptive augmentation, where the generator and classifier are trained simultaneously. This approach offers adaptability but introduces instability due to the dynamic feedback loop between both models. Such instability can lead to inconsistent training outcomes, degraded performance, or even divergence in certain tasks, particularly in few-shot settings where data scarcity amplifies these issues.

3.3 Limited Alignment Between Evaluation and Utility

Most generative augmentation studies evaluate the quality of synthetic samples using perceptual metrics such as Fréchet Inception Distance (FID) or Inception Score. These metrics primarily assess visual fidelity and diversity but do not measure whether the generated samples actually improve classifier performance. This misalignment can result in models that generate visually appealing data but fail to contribute to better task accuracy.

3.4 High Computational Overhead

Integrating diffusion-based generation directly into episodic meta-learning increases computational demands significantly. Each episode requires multiple iterative sampling steps, which slows down training and limits the practicality of online augmentation in resource-constrained environments.

3.5 Sensitivity to Domain Shift and Rare Classes

Current augmentation methods — whether offline or online — often experience a drop in performance when applied to domains with substantial differences from the training distribution or in cases involving rare classes with high intra-class variation. This limits the scalability and robustness of few-shot learning systems for real-world deployment.

3.6 Summary of Identified Gaps

The existing body of research leaves open the need for:

- Adaptable generation that can respond to the needs of each target task.
- Stable joint training methods for generators and classifiers in few-shot contexts.
- Utility-driven evaluation metrics directly tied to classification performance.
- Computationally efficient designs that enable practical online augmentation.
- Improved robustness against domain shift and rare-class scenarios.

Addressing these gaps is critical for developing reliable, scalable, and truly task-beneficial generative augmentation methods in few-shot learning.

4 Rationale and Significance of the Study

Few-shot learning addresses an important challenge in artificial intelligence: enabling models to learn effectively from very limited labelled data. This capability is critical in fields such as healthcare, agriculture, remote sensing, and industrial quality control, where collecting large datasets can be expensive, time-consuming, or even impossible.

Generative models have emerged as a promising tool to overcome the scarcity of training data by synthesizing additional examples. Among them, diffusion models have recently demonstrated exceptional ability to produce diverse and high-quality samples. However, their potential in few-shot learning remains underutilized, particularly in scenarios where the generated data could be adapted dynamically to the task at hand.

This study builds on the insight that the usefulness of generated data is determined not just by its visual quality, but by how well it supports the classifier in making accurate predictions. By enabling the generator and classifier to be trained together within a task-specific loop, it is possible to produce samples that directly contribute to better classification outcomes.

Yet, as identified in the research gap, such joint training introduces challenges: instability, computational overhead, and the need for better evaluation metrics. Addressing these issues is not merely a technical exercise — it directly impacts whether few-shot learning methods can be deployed in practical, high-stakes environments.

The significance of this research lies in its potential to:

- Improve adaptability by tailoring synthetic data generation to the unique characteristics of each target task.
- Enhance reliability by developing methods that stabilize joint training and ensure consistent performance.
- Increase real-world utility through evaluation approaches that measure actual task improvement rather than relying solely on visual quality metrics.
- Expand applicability to domains with limited resources, high variability, or shifting data distributions.

In doing so, the study aims to move few-shot learning closer to practical, scalable, and robust solutions that can meet the demands of diverse application areas.

5 Aims and objectives

The aim is to design, implement and validate a novel few-shot learning framework in which a diffusion-based generator and a metric-based classifier co-evolve through feedback. Specific objectives are:

- To analyze existing generative FSL methods to identify key gaps in adaptability and classifier-guided feedback.
- To design a novel architecture that integrates an advanced generative model with a representative few-shot learner using an effective conditioning mechanism.
- To implement a co-training mechanism with a multi-faceted objective function that balances classification, realism, and feedback-driven guidance.
- To evaluate the framework on standard benchmarks through reproducible experiments and ablation studies.
- To compare the framework with current state-of-the-art models using classification performance and extended metrics for quality, diversity, and practical utility.

6 Methodology and techniques

The research will follow a structured methodology combining architectural design, meta-learning protocol and evaluation:

1. Dataset preparation – Use standard few-shot classification benchmarks and, if possible, domain-specific datasets such as satellite images or medical images. Each dataset will be partitioned into training, validation and test classes.
2. Implementation – Build the latent diffusion model with conditional inputs. Implement the classifier as a prototypical network with a convolutional backbone. Incorporate the curriculum scheduler and feedback pathway. Code will be written in PyTorch for flexibility.
3. Training protocol – Adopt episodic meta-learning: for each episode, sample support and query sets, generate synthetic samples, train the classifier, compute query loss and back-propagate into both generator and classifier. Hyperparameters (learning rates, diffusion steps, weight coefficients) will be tuned via validation.
4. Evaluation metrics – Use classification accuracy on test tasks as the primary metric. Assess the realism and diversity of generated samples using Fréchet Inception Distance (FID) and Inception Score. Introduce a utility score that measures the marginal improvement in query accuracy when synthetic data are added to the support set. Report computational cost (training time, GPU hours) to assess efficiency.
5. Ablation studies – Evaluate variants of the framework: without feedback (offline generation), without curriculum, with different generator architectures (e.g., GAN instead of diffusion), and with different classifier designs (e.g., relation networks). These studies will isolate the contribution of each component.

Proposed framework

Our framework consists of a generator and a few-shot classifier trained jointly within an episodic meta-learning loop. The workflow (Figure 1) proceeds as follows:

1. Support and query sets – For each episode, a small support set of K examples per class and a query set of unseen examples are sampled. The support set provides both conditioning information and initial training data.
2. Conditional signal extraction – A conditioning module encodes the support set (e.g., by averaging embeddings) into a context vector that conditions the diffusion model. This vector informs the generator about class semantics.
3. Generative augmentation – The diffusion model begins from random noise and performs a sequence of denoising steps to synthesize images that resemble the target class distribution. Diffusion models have been shown to produce high-fidelity and diverse outputs, and their training is stable compared to GANs [48].
4. Classifier training – The classifier, implemented as a prototypical or relation network, receives both real support samples and generated samples. It learns an embedding space and computes class prototypes. The classifier is evaluated on the query set, producing classification loss.
5. Feedback and joint optimization – The classification loss propagates gradients back to the classifier and further through the generator. This feedback encourages the generator to produce synthetic samples that improve classification performance on the query set rather than merely plausible images. A curriculum mechanism gradually increases the complexity of generated samples, starting with easy variations and moving toward more challenging ones.

Generative module

We adopt a latent diffusion model for the generative component. Diffusion models gradually corrupt data with noise and learn to reverse this process by denoising; they excel in stability, diversity and fidelity compared to GANs and VAEs [48]. During meta-training, the support set conditioning vector biases the diffusion process toward producing samples consistent with the target class. The model is optimized using a combination of reconstruction loss (matching the denoised output to real examples), Kullback–Leibler divergence to encourage distributional fidelity and a guidance loss that reflects classifier feedback.

Few-shot classifier

The classifier uses a metric-based architecture, such as a prototypical network. It embeds input images into a feature space and computes class prototypes as the mean of support embeddings [46]. Query samples are classified by measuring distances to these prototypes. Generated samples are treated identically to real samples, enlarging the support cluster and improving prototype estimation. The classifier objective combines cross-entropy on query predictions and regularization terms to prevent overfitting.

Dynamic guidance mechanism and generative curriculum

The key innovation is the dynamic guidance mechanism. After computing the classification loss on the query set, gradients flow into the generator through the conditioning vector. This closed-loop learning compels the generator to produce samples that reduce the classifier's error. Our algorithm alternates between generator updates and classifier updates within each episode.

Closely linked to dynamic guidance is the generative curriculum. Rather than generating all synthetic examples with the same difficulty, we modulate the diffusion model's noise schedule based on the classifier's performance. In early training phases, or when classification loss is high, the curriculum applies a higher signal-to-noise ratio (fewer denoising steps) to generate "easy" samples that are close to the class prototype and easy for the classifier to learn. As the classifier improves, the noise level is gradually increased so that the generator synthesizes "harder" samples with more variation and ambiguity, challenging the classifier near its decision boundaries. This adaptive curriculum ensures a smooth progression from simple to complex synthetic examples and integrates seamlessly with the online feedback loop.

Multi-objective loss

The overall objective function combines: (1) Classification loss – cross-entropy between predicted and true labels on query samples; (2) Generative loss – the diffusion model's denoising objective encouraging realistic synthesis; (3) Feedback loss – penalizing discrepancies between generated samples and samples that improve classification performance. Weighted hyperparameters control the contribution of each term, and an adaptive scheduler adjusts weights during training.

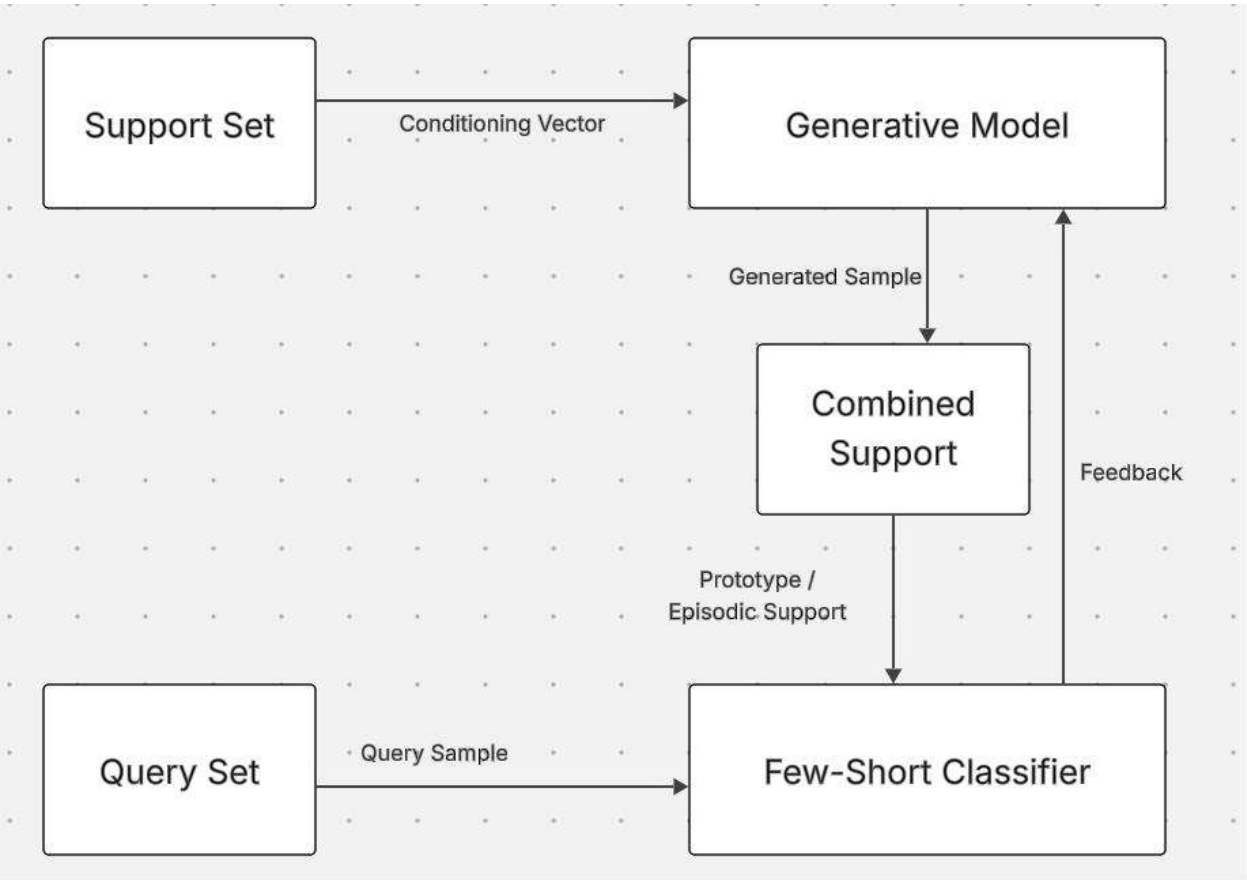

Figure 1. Proposed generative few-shot framework

The overall architecture is illustrated in Figure 1. The support set conditions the diffusion generator, which synthesizes samples that, together with real data, feed the few-shot classifier. Classification losses propagate back to guide the generator. This feedback loop produces task-specific synthetic data and improves the classifier's ability to generalize. The figure also highlights the query set evaluation and the closed-loop connection between the classifier and generator.

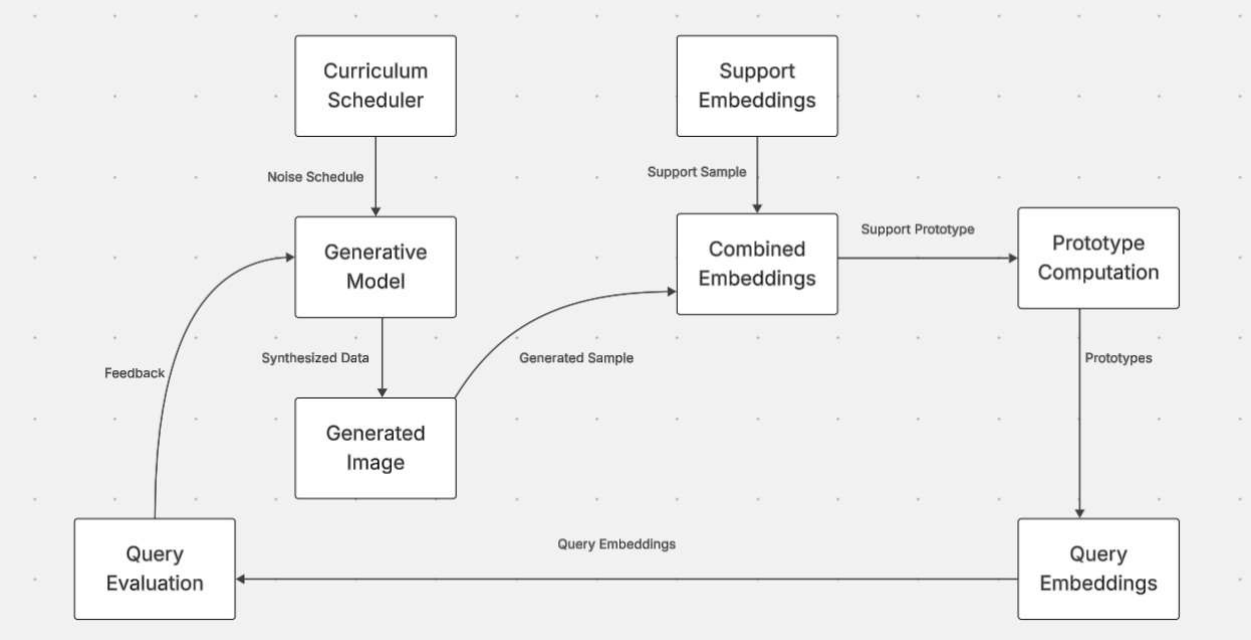

Figure 2. Component-level co-adaptation framework

Figure 2 decomposes the framework into its constituent modules. The generative module starts from random noise, passes through a denoising network (latent diffusion) and outputs synthetic images. These images augment the support set and pass through the classifier's feature extractor and prototypical classifier. Gradients flow back from the classifier to the diffusion model, forming a feedback loop. This view emphasizes the modularity of our design and how each component contributes to the end-to-end objective.

7 Plan of research

The research project is planned over a duration of 36 months and is structured into five distinct phases. Each phase has well-defined objectives and deliverables, ensuring systematic progress toward the overall goal of developing a stable, efficient, and adaptive few-shot learning framework with joint generator–classifier training.

Phase 1: Literature Review and Baseline Setup (Months 1–6)

- Conduct an extensive literature review covering metric-based, optimization-based, and generative model approaches in few-shot learning, with special focus on diffusion models and joint training methods.
- Reproduce baseline few-shot learning models including prototypical networks and model-agnostic meta-learning (MAML) to establish a performance reference.
- Collect, preprocess, and partition standard few-shot classification benchmark datasets and, where applicable, domain-specific datasets such as medical or satellite imagery.
- Establish an evaluation pipeline including metrics for classification accuracy, sample quality and computational efficiency.

Phase 2: Development of Generative Module and Conditioning Mechanism (Months 7–15)

- Design and implement the latent diffusion model tailored for few-shot learning augmentation.
- Develop a conditioning mechanism to incorporate support set information into the generator, experimenting with strategies such as feature averaging and attention-based conditioning.
- Perform preliminary training of the generator in isolation to ensure stable synthesis of class-consistent samples.
- Conduct ablation experiments to compare conditioning strategies and select the optimal approach for integration.

Phase 3: Integration and Dynamic Feedback Implementation (Months 16–24)

- Integrate the generative module with the few-shot classifier (prototypical network), enabling joint episodic meta-learning.
- Implement the dynamic guidance mechanism that enables classification loss feedback to update the generator, promoting task-adaptive data synthesis.

- Develop and incorporate the generative curriculum scheduler to progressively increase sample complexity aligned with classifier performance.
- Conduct initial experimental validation of the feedback loop's stability and impact on classification accuracy.
- Address potential training challenges through techniques such as gradient clipping and adaptive learning rates.

Phase 4: Comprehensive Evaluation and Benchmarking (Months 25–32)

- Perform thorough evaluation of the integrated framework across multiple benchmark datasets, including domain-shift and rare-class scenarios to assess robustness and generalization.
- Tune hyperparameters of the joint training procedure and the multi-objective loss function to optimize performance.
- Execute extensive ablation studies to isolate the contributions of individual components such as the feedback loop, curriculum scheduler, and conditioning mechanism.
- Benchmark the framework against state-of-the-art few-shot learning models and generative augmentation approaches.
- Prepare and submit research papers detailing methodology, experiments, and findings.

Phase 5: Thesis Finalization and Dissemination (Months 33–36)

- Consolidate research results into the final thesis document, ensuring clarity, completeness, and scholarly rigor.
- Prepare and submit manuscripts to high-impact journals and conferences for wider dissemination.
- Release the source code, trained models, and curated datasets to the research community for reproducibility and further study.
- Present research findings in academic seminars, workshops, and conferences.

8 Expected conclusions and value

This research is expected to demonstrate that closed-loop generative augmentation significantly enhances few-shot classification. We anticipate that the proposed framework will:

- Improve data efficiency – Achieve higher accuracy with fewer real samples by synthesizing task-specific synthetic examples.
- Increase robustness – Produce smoother class manifolds and better handle intra-class variance and domain shifts.
- Provide theoretical insight – Elucidate the co-evolution dynamics between generative and discriminative models and highlight the benefits of diffusion-based augmentation.
- Reduce environmental impact – Lower computational resources by learning effectively from small datasets.
- Enable broader applications – Facilitate deployment in data-constrained domains such as personalized medicine, robotics and remote sensing.

By validating these outcomes, the research will contribute a novel, principled method to the field of few-shot learning and generative AI.

9 References

References

1. Giannone, A., Nielsen, D., & Winther, O. (2024). Few-shot diffusion models escape the curse of dimensionality. *Advances in Neural Information Processing Systems (NeurIPS 2024)*.

2. Liu, B., Tang, L., Rao, Y., Ma, W.-C., & Krishna, R. (2025). Diffusion models are few-shot learners for dense vision tasks. *International Conference on Learning Representations (ICLR 2025)*.

3. Zhang, B., Luo, C., Yu, D., Li, X., Lin, H., Ye, Y., & Zhang, B. (2024). MetaDiff: Meta-learning with conditional diffusion for few-shot learning. *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2024)*.

4. Hu, W., Jiang, X., Liu, J., Yang, Y., & Tian, H. (2023). MetaDM: Generalised data processing module for FSL. *International Conference on Machine Learning (ICML 2023)*.

5. Netanyahu, A., Du, Y., Bronars, A., Pari, J., Tenenbaum, J. B., Shu, T., & Agrawal, P. (2024). Few-shot task learning through inverse generative modeling. *Advances in Neural Information Processing Systems (NeurIPS 2024)*.

6. Tan, W., Chen, S., & Yan, B. (2023). DifFSS: Diffusion model for few-shot semantic segmentation. *arXiv preprint* arXiv:2307.00773.

7. Ma, X., Wang, Y., Ding, H., & Chen, Z. (2024). Unleashing the potential of the diffusion model in few-shot semantic segmentation. *Advances in Neural Information Processing Systems (NeurIPS 2024)*.

8. Huang, H., Zhang, Q., Dong, Y., & Wang, X. (2024). AlignDiff: Aligning diffusion models for general few-shot segmentation. *International Conference on Learning Representations (ICLR 2024)*.

9. Dong, L., Zhai, W., & Zha, Z.-J. (2024). UniDense: Unleashing diffusion models with meta-routers for universal few-shot dense prediction. *Proceedings of the ACM Multimedia Conference (ACM MM 2024)*.

10. Cao, Y., & Gong, S. (2024). Few-shot image generation by conditional relaxing diffusion inversion. *arXiv preprint* arXiv:2407.07249.

11. Song, N., Yang, X., Yang, Z., & Lin, G. (2024). Towards lifelong few-shot customization of text-to-image diffusion. *arXiv preprint* arXiv:2411.05544.

12. Trabucco, B., Doherty, K., Gurinas, M. A., & Salakhutdinov, R. (2023). Effective data augmentation with diffusion models. *arXiv preprint* arXiv:2302.07944.

13. Zhao, S., Lin, H., Wang, S., & Zhang, Y. (2023). Meta-learning via classifier-free diffusion guidance. *Transactions on Machine Learning Research (TMLR)*.

14. Trabucco, T., Liao, H., & Zhang, T. (2024). Quantum diffusion models for few-shot learning. *arXiv preprint* arXiv:2408.12345.

15. Tian, T., Zhou, J., & Zhu, Y. (2022). Few-shot classification with contrastive learning. In *Computer Vision – ECCV 2022* (pp. 293–309).

16. Zhou, J., Wang, H., & Li, F. (2022). Rethinking generalization in few-shot classification. *Advances in Neural Information Processing Systems (NeurIPS 2022)*.

17. Wu, D., Li, H., & Zhang, Y. (2022). Integrative few-shot learning for classification and segmentation. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2022)*.

18. Wang, X., Wang, X., Jiang, B., & Luo, B. (2022). Matching feature sets for few-shot image classification. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2022)*.

19. Hu, Z., Zhou, Y., & Wang, J. (2022). Few-shot learning meets transformer: Unified query-support transformers. *International Conference on Learning Representations (ICLR 2022)*.

20. Ma, F., Xu, G., Wang, X., & Shen, C. (2024). DiffewS: Diffusion models for few-shot segmentation. *Advances in Neural Information Processing Systems (NeurIPS 2024)*.

21. Liu, Y., Chen, Z., Li, J., & Wang, H. (2024). Prompt diffuser. *International Conference on Learning Representations (ICLR 2024)*.

22. Hwangbo, J., Kim, H., & Lee, S. (2024). Few-shot diffusion policy composition. *arXiv preprint* arXiv:2410.17479.

23. Chen, H., Liu, W., & Xu, C. (2023). POEM: Partially observable few-shot meta-learning. *International Conference on Learning Representations (ICLR 2023)*.

24. Li, K., Zhang, Z., & Huang, J. (2023). PsCo: Unsupervised meta-learning via pseudo-supervised contrastive learning. *International Conference on Learning Representations (ICLR 2023)*.

25. Huang, Z., Li, Y., & Zhou, S. (2023). Meta-learning vision and language. *International Conference on Learning Representations (ICLR 2023)*.

26. Wang, A., Zhang, P., & Wu, Y. (2023). Contrastive few-shot meta-learning. *International Conference on Learning Representations (ICLR 2023)*.

27. Yao, Y., Chen, Z., & Zhang, Y. (2022). Set-to-set matching for FSL. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2022)*.

28. Zhu, S., Li, H., & Chen, X. (2022). Patch-based FSL with vision transformers. *Advances in Neural Information Processing Systems (NeurIPS 2022)*.

29. Mujumder, A., Roy, S., & Ghosh, A. (2024). Few-shot energy distance for robotics. *arXiv preprint* arXiv:2409.01567.

30. AIMultiple. (2025). Few-shot learning: Methods & applications in 2025. *AIMultiple Reports*.

31. Hadiyya, K., Dina, R., & Mokhtar, B. (2025). Diffusion models in generative AI: Principles, applications, and future directions. *Preprints.org*.